

IN THE U.S. PATENT AND TRADEMARK OFFICE

In re application of

Satoshi HOSOKAWA

Conf.

Application No. NEW NON-PROVISIONAL

Group

Filed March 11, 2004

Examiner

APPARATUS AND METHOD FOR IMAGE PROCESSING, AND IMAGE  
PROCESSING SYSTEM USING THE SAME AND PROGRAM

CLAIM TO PRIORITY

Assistant Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

March 11, 2004

Sir:

Applicant(s) herewith claim(s) the benefit of the  
priority filing date of the following application(s) for the  
above-entitled U.S. application under the provisions of 35  
U.S.C. § 119 and 37 C.F.R. § 1.55:

<u>Country</u>	<u>Application No.</u>	<u>Filed</u>
JAPAN	2003-084034	March 26, 2003

Certified copy(ies) of the above-noted application(s)  
is(are) attached hereto.

Respectfully submitted,

YOUNG & THOMPSON



---

Benoit Castel, Reg. No. 35,041  
745 South 23<sup>rd</sup> Street  
Arlington, VA 22202  
Telephone (703) 521-2297  
Telefax (703) 685-0573  
703) 979-4709

BC/ia

Attachment(s): 1 Certified Copy(ies)

日本国特許庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2003年 3月26日  
Date of Application:

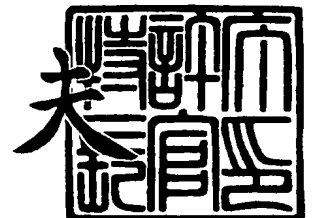
出願番号 特願2003-084034  
Application Number:  
[ST. 10/C]: [JP2003-084034]

出願人 日本電気株式会社  
Applicant(s):

2004年 2月 3日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



出証番号 出証特2004-3005566



【書類名】 特許願

【整理番号】 53211024

【提出日】 平成15年 3月26日

【あて先】 特許庁長官殿

【国際特許分類】 H04N 5/91

【発明者】

    【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

    【氏名】 細川 知志

【特許出願人】

    【識別番号】 000004237

    【氏名又は名称】 日本電気株式会社

【代理人】

    【識別番号】 100088812

    【弁理士】

    【氏名又は名称】 ▲柳▼川 信

【手数料の表示】

    【予納台帳番号】 030982

    【納付金額】 21,000円

【提出物件の目録】

    【物件名】 明細書 1

    【物件名】 図面 1

    【物件名】 要約書 1

    【包括委任状番号】 9001833

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像処理装置及びその方法並びにそれを用いた画像処理システム及びプログラム

【特許請求の範囲】

【請求項 1】 静止画像データを所定の動画像符号化方式により符号化する動画符号化手段と、

前記動画符号化手段の出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工手段と、

前記データ加工手段の出力と前記音声データとを多重化して出力する多重化手段と、

を含むことを特徴とする画像処理装置。

【請求項 2】 前記音声データを所定の音声符号化方式により符号化する音声符号化手段を、更に含むことを特徴とする請求項 1 記載の画像処理装置。

【請求項 3】 前記データ加工手段は、前記音声符号化手段から得られる前記音声データの再生時間に関する情報の供給を受けて、前記時間伸長情報を得ることを特徴とする請求項 1 記載の画像処理装置。

【請求項 4】 前記データ加工手段は、

前記動画符号化手段の出力から前記静止画データに相当する 1 フレーム分のデータを抽出し、当該データの後に、前記音声データの再生時間に相当する時間だけ、所定周期で前記時間伸長情報として無効フレームを付加する第一の手段を有することを特徴とする請求項 1 または 2 記載の画像処理装置。

【請求項 5】 前記データ加工手段は、

動画符号化手段の出力から前記静止画データに相当する 1 フレーム分のデータを抽出し、前記音声データの再生時間に相当する時間後に前記時間伸長情報として無効フレームを付加する第二の手段を有することを特徴とする請求項 1 または 2 記載の画像処理装置。

【請求項 6】 前記データ加工手段は、

動画符号化手段の出力から前記静止画データに相当する 1 フレーム分のデータを抽出し、この抽出データに前記時間伸長情報を付加する第三の手段を有することを特徴とする請求項 1 または 2 記載の画像処理装置。

【請求項 7】 前記データ加工手段は、

前記動画符号化手段の出力から前記静止画データに相当する 1 フレーム分のデータを抽出する抽出手段と、

当該データの後に、前記音声データの再生時間に相当する時間だけ、所定周期で前記時間伸長情報として無効フレームを付加する第一の手段と、

前記音声データの再生時間に相当する時間後に前記時間伸長情報として無効フレームを付加する第二の手段と、

この抽出データに前記時間伸長情報を付加する第三の手段と、

これ等第一～第三の手段を選択する選択手段と、

を含むことを特徴とする請求項 1 または 2 記載の画像処理装置。

【請求項 8】 前記データ加工手段は、

前記第一、第二、または第三の手段の処理終了後に、更に所定回数前記第一、第二、または第三の手段の処理を続行制御する手段を、更に有することを特徴とする請求項 4～7 いずれか記載の画像処理装置。

【請求項 9】 前記静止画像データを生成するカメラと、前記音声データを生成するマイクとを、更に含むことを特徴とする請求項 1～8 いずれか記載の画像処理装置。

【請求項 10】 携帯電話機であることを特徴とする請求項 9 記載の画像処理装置。

【請求項 11】 音声データが付加された動画データから所望の静止画像データとそれに関連する音声データとを抽出する手段を、更に含み、この抽出静止画像データを前記加工手段へ供給するようにしたことを特徴とする請求項 1～7 いずれか記載の画像処理装置。

【請求項 12】 静止画像データを所定の動画符号化方式により符号化する動画符号化ステップと、

この符号化出力に、前記静止画像データに関連する音声データの再生時間に等

しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工ステップと、

このデータ加工後の出力と前記音声データとを多重化して出力する多重化ステップと、

を含むことを特徴とする画像処理方法。

【請求項 13】 前記音声データを所定の音声符号化方式により符号化する音声符号化ステップを、更に含むことを特徴とする請求項 12 記載の画像処理方法。

【請求項 14】 前記データ加工ステップは、前記音声符号化ステップから得られる前記音声データの再生時間に関する情報の供給を受けて、前記時間伸長情報を得ることを特徴とする請求項 12 記載の画像処理方法。

【請求項 15】 前記データ加工ステップは、前記動画符号化ステップの出力から前記静止画データに相当する 1 フレーム分のデータを抽出し、当該データの後に、前記音声データの再生時間に相当する時間だけ、所定周期で前記時間伸長情報として無効フレームを付加する第一のステップを有することを特徴とする請求項 12 または 13 記載の画像処理方法。

【請求項 16】 前記データ加工ステップは、動画符号化ステップの出力から前記静止画データに相当する 1 フレーム分のデータを抽出し、前記音声データの再生時間に相当する時間後に前記時間伸長情報として無効フレームを付加する第二のステップを有することを特徴とする請求項 12 または 13 記載の画像処理方法。

【請求項 17】 前記データ加工ステップは、動画符号化ステップの出力から前記静止画データに相当する 1 フレーム分のデータを抽出し、この抽出データに前記時間伸長情報を付加する第三のステップを有することを特徴とする請求項 12 または 13 記載の画像処理方法。

【請求項 18】 前記データ加工ステップは、前記動画符号化ステップの出力から前記静止画データに相当する 1 フレーム分のデータを抽出する抽出ステップと、

当該データの後に、前記音声データの再生時間に相当する時間だけ、所定周期

で前記時間伸長情報として無効フレームを付加する第一のステップと、

前記音声データの再生時間に相当する時間後に前記時間伸長情報として無効フレームを付加する第二のステップと、

この抽出データに前記時間伸長情報を付加する第三のステップと、

これ等第一～第三のステップを選択する選択ステップと、  
を含むことを特徴とする請求項 1 2 または 1 3 記載の画像処理方法。

【請求項 1 9】 前記データ加工ステップは、

前記第一、第二、または第三のステップの処理終了後に、更に所定回数前記第一、第二、または第三のステップを続行制御するステップを、更に有することを特徴とする請求項 1 5 ～ 1 8 いずれか記載の画像処理方法。

【請求項 2 0】 前記静止画像データをカメラから生成するステップと、前記音声データをマイクから生成するステップとを、更に含むことを特徴とする請求項 1 2 ～ 1 9 いずれか記載の画像処理方法。

【請求項 2 1】 携帯電話機であることを特徴とする請求項 2 0 記載の画像処理方法。

【請求項 2 2】 音声データが付加された動画像データから所望の静止画像データとそれに関連する音声データとを抽出するステップを、更に含み、この抽出静止画像データを前記加工ステップへ供給するようにしたことを特徴とする請求項 1 2 ～ 1 8 いずれか記載の画像処理方法。

【請求項 2 3】 静止画像データを所定の動画像符号化方式により符号化する動画符号化手段と、前記動画符号化手段の出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画符号化方式に応じた形式で付加するデータ加工手段と、前記データ加工手段の出力と前記音声データとを多重化して出力する多重化手段とを含む画像処理装置と、

前記画像処理装置からの多重化出力を受信して、前記データ加工手段により加工されたデータを前記動画符号化方式に対応する復号化方式により復号化する動画復号化手段と、前記多重化出力の音声データと前記動画復号化手段と同時再生する手段とを含む再生装置と、

を備えてなることを特徴とする画像処理システム。

【請求項 2 4】 静止画像データを所定の動画像符号化方式により符号化する動画像符号化手段と、前記動画像符号化手段の出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工手段と、前記データ加工手段の出力と前記音声データとを多重化して出力する多重化手段とを含む画像処理装置からの多重化出力を受信して再生する再生装置であって、

前記データ加工手段により加工されたデータを前記動画像符号化方式に対応する復号化方式により復号化する動画像復号化手段と、

前記多重化出力の音声データと前記動画像復号化手段と同時再生する手段と、を含むことを特徴とする再生装置。

【請求項 2 5】 画像処理装置の画像処理動作をコンピュータにより実行させるためのプログラムであって、

静止画像データを所定の動画像符号化方式により符号化する動画像符号化ステップと、

この符号化出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工ステップと、

このデータ加工後の出力と前記音声データとを多重化して出力する多重化ステップと、を含むコンピュータ読取り可能なプログラム。

#### 【発明の詳細な説明】

##### 【0 0 0 1】

#### 【発明の属する技術分野】

本発明は画像処理装置及びその方法並びにそれを用いた画像処理システム及びプログラムに関し、特に静止画像と音声とを同時に再生処理するためのファイルを作成する画像処理方式に関するものである。

##### 【0 0 0 2】

#### 【従来の技術】



静止画像と共に、それに関連する音声と同時に再生処理することが必要な場合がある。例えば、携帯端末である携帯電話機には、画像を撮影するカメラ付きのものがあ、このカメラにより撮影した静止画像に対して、それに関連した説明などの音声を同時に付加して、相手先の携帯電話機やP C（パーソナルコンピュータ）へ送信することが要求されるようになってきている。また、逆に、P Cから、静止画像にその音声を付加して、携帯電話機や他のP Cに対して送信することも考えられる。

#### 【0003】

このような、静止画像とそれに関連する音声情報とを得るための技術としては、特許文献1にあるように、静止画像としての1フレーム分の映像信号と、この1フレーム分の映像信号の内容を説明するための音声信号を時間軸圧縮した信号とを周波数多重化して、記録媒体の同一領域に記録する。そして、再生の際に、時間軸伸長して元の音声信号に戻し、映像信号の繰り返し再生中に当該音声信号を出力するようにして、静止画像の説明を対応する静止画像と同時に出力することが可能となっている。

#### 【0004】

また、他の方法として、静止画像とこれに関連する音声データとを、特定の方式で単一のファイルにまとめ、その特定の方式に対応した再生装置で再生する第一の方法や、静止画像とこれに関連する音声データとを、互いに別々に独立して扱った上で、再生装置にこれ等両者を同時に再生するための専用の機能を設ける第二の方法が考えられる。

#### 【0005】

##### 【特許文献1】

特開昭64-86364号公報（第3頁、第1図）

#### 【0006】

##### 【発明が解決しようとする課題】

上述した特許文献1に開示の技術では、再生の際に、音声信号の再生時間に相当する時間長の静止画像を、繰り返し再生するようになっていたために、静止画像の繰り返しの比例して情報量が増大することになる。そのために、上述した携

携帯電話機やPCの再生側において、再生のための情報量が多くなって、処理能力が問題となる。

【0007】

また、上述した第一の方法では、特定の方式に対応した専用の再生装置が必要になる。このため、その方式に対応していない再生装置（過去に存在した再生装置など）では再生することができず、よって、現在の携帯電話機やPCなどでは対応が不可能となる。

【0008】

また、上述した第二の方法では、静止画像データと音声データとは、普及している形式を使用することができるために、静止画像と音声とを、それぞれ再生することは可能であるが、これ等を同時再生に関する考慮がなされていない再生装置では、同時再生ができないばかりか、静止画と音声がばらばらに取り扱われてしまう可能性がある。よって、この第二の方法も現在の携帯電話機やPCなどには対応が不可能である。

【0009】

本発明の目的は、特定の形式や専用の機能をもたない再生装置においても、静止画像とそれに関連する音声との同時再生を可能とするようにした画像処理装置及びその方法並びにそれを用いた画像処理システム及びプログラムを提供することである。

【0010】

本発明の他の目的は、情報量の増大なく、静止画像とそれに関連する音声との同時再生を可能とするようにした画像処理装置及びその方法並びにそれを用いた画像処理システム及びプログラムを提供することである。

【0011】

【課題を解決するための手段】

本発明による画像処理装置は、静止画像データを所定の動画像符号化方式により符号化する動画像符号化手段と、前記動画像符号化手段の出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工手段

と、前記データ加工手段の出力と前記音声データとを多重化して出力する多重化手段とを含むことを特徴とする。

#### 【0012】

本発明による画像処理方法は、静止画像データを所定の動画像符号化方式により符号化する動画像符号化ステップと、この符号化出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工ステップと、このデータ加工後の出力と前記音声データとを多重化して出力する多重化ステップとを含むことを特徴とする。

#### 【0013】

本発明による画像処理システムは、静止画像データを所定の動画像符号化方式により符号化する動画像符号化手段、前記動画像符号化手段の出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工手段、前記データ加工手段の出力と前記音声データとを多重化して出力する多重化手段を含む画像処理装置と、前記画像処理装置からの多重化出力を受信して、前記データ加工手段により加工されたデータを前記動画像符号化方式に対応する復号化方式により復号化する動画像復号化手段、前記多重化出力の音声データと前記動画像復号化手段と同時再生する手段を含む再生装置とを備えてなることを特徴とする。

#### 【0014】

本発明による再生装置は、静止画像データを所定の動画像符号化方式により符号化する動画像符号化手段と、前記動画像符号化手段の出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工手段と、前記データ加工手段の出力と前記音声データとを多重化して出力する多重化手段とを含む画像処理装置からの多重化出力を受信して再生する再生装置であって、前記データ加工手段により加工されたデータを前記動画像符号化方式に対応する復号化方式により復号化する動画像復号化手段と、前記多重化出力の音声データと

前記動画復号化手段と同時再生する手段とを含むことを特徴とする。

【0015】

本発明によるプログラムは、画像処理装置の画像処理動作をコンピュータにより実行させるためのプログラムであって、静止画像データを所定の動画像符号化方式により符号化する動画像符号化ステップと、この符号化出力に、前記静止画像データに関連する音声データの再生時間に等しい時間だけ再生時間を伸長する旨の時間伸長情報を、前記動画像符号化方式に応じた形式で付加するデータ加工ステップと、このデータ加工後の出力と前記音声データとを多重化して出力する多重化ステップとを含むことを特徴とする。

【0016】

本発明の作用を述べる。所定の動画像符号化方式を用いて符号化された静止画像データに対して、動画再生装置において静止画像の再生を所定時間継続（時間伸長）させるための情報を当該動画像符号化方式に応じた形式で付与してファイルを作成する。この場合、静止画像データを単に符号化するだけでは、これを動画再生装置で再生する場合に、静止画像の表示は極く短時間で終了してしまう。従って、単に符号化された静止画像データとこれに対応する音声データとを多重化することにより得られた動画ファイルを動画再生装置で再生する場合には、静止画像の表示と音声の再生とを同時に行なうことができない。すなわち、音声を再生している間、静止画像の表示を継続させることができない。

【0017】

そこで、本発明では、上述したように、所定の動画像符号化方式で符号化された静止画像に対して、音声データの再生時間に相当する時間だけ伸長する旨の時間伸長情報を付加することにより、これを動画再生装置で再生する場合、符号化された静止画像データに付与された時間伸長情報により、静止画像の表示を所定時間継続させることができる。

【0018】

従って、こうして得られた静止画像データとこれに対応する音声データとを多重化することにより得られた動画ファイルを、動画再生装置で再生する場合、音声を再生している間、静止画像の表示を継続して行うことが可能となる。

**【0019】****【発明の実施の形態】**

以下に、本発明の実施例につき図面を参照しつつ詳細に説明する。図1は本発明の一実施例のシステム構成を示す図である。図1を参照すると、記録装置1と再生装置11とがネットワーク10を介して接続されており、記録装置1はPCや携帯型情報処理端末であり、また再生装置11も、同じく、PCや携帯型情報処理端末であるものとする。

**【0020】**

記録装置1は静止画像ファイル2と音声ファイル3とを備えており、また、静止画像ファイル2からの静止画像データを入力として、動画データの生成を行うための動画エンコーダ4と、音声ファイル3からの音声データを入力として動画データに対応した音声データを生成する音声エンコーダ5とを備えている。更に、動画エンコーダ4の出力と音声エンコーダ5からの音声再生時間情報とを入力として、静止画像をこの音声再生時間情報により示される時間長の映像データとして加工処理するためのデータ加工部6を備えている。

**【0021】**

記録装置1は、更に、データ加工部6と音声エンコーダ5との出力、すなわち映像と音声との多重化を行って、動画ファイルを生成する機能を有している。この記録装置1におけるデータ加工部6以外の部分は、映像ファイルと音声ファイルとを多重化して、動画ファイルを生成する機能を有する一般の映像記録装置と同一の構成である。そして、マルチプレクサ7の多重化出力がネットワーク10を介して再生装置11へ送信されることになる。

**【0022】**

この再生装置11は、ネットワーク10を介して送られてきた動画ファイルの再生を行うためのものであり、そのために、映像と音声との多重化を解除して、これ等を別々に導出するデマルチプレクサ12を備えている。また、映像と音声とをそれぞれ入力として、記録装置1側の動画エンコーダ4、音声エンコーダ5の各エンコード方式に対応したデコード方式を有する動画デコーダ13、音声デコーダ14を備えており、これ等デコード出力は表示部15、スピーカ16へ供

給されて静止画とそれに対応する説明用の音声とが、同時に再生されることになる。

#### 【0023】

再生装置 11 は、以上の機能により動画ファイルの再生を行うことが可能であるが、静止画像と音声とを同時再生するための特別な専用の機能を備えているわけではなく、一般的な動画再生装置である。

#### 【0024】

本発明における動画ファイル作成の過程を図 2 に示している。図 2 において、図 1 と同等部分は同一符号にて示している。図 2 を参照すると、記録装置 1 においては、データ加工部 6 によって、静止画像が映像としてエンコードされて動画ファイル D として作成される。これにより、映像が 1 フレームだけ記録された動画ファイルが作成されることになる。すなわち、動画エンコーダ 4 に静止画ファイル 2 の静止画像データが入力され、1 フレーム分のエンコードが行われる。この時点で生成される映像データ A は 1 フレーム分のみであるために、非常に短い再生時間となっている（1 フレームの時間は十数分の一秒）。この 1 フレーム分のデータ A はデータ加工部 6 へ入力されて、1 フレーム分のデータを音声ファイル 3 の長さと同じになるように再生時間の伸長擬装をなす加工処理が行われる。更に、この擬装処理された映像データ B と、音声ファイル 3 を音声エンコーダ 5 によりエンコードして得られた音声データ C とが、マルチプレクサ 7 へ入力されて動画ファイル D として生成されるのである。

#### 【0025】

上述の方法により、記録装置 1 にて生成された動画ファイルを、再生装置 11 において再生すると、記録装置 1 の表示部 15 には、動画デコーダ 13 によりデコードされた静止画像の内容が表示され、スピーカ 16 からは、音声デコーダ 14 によりデコードされた音声再生される。このとき、映像データは 1 フレーム分のデータのみであるが、その長さは動画ファイルの長さ（すなわち、元の音声ファイル 3 の再生時間に相当する長さ）と同一であるように時間長（時間伸長）擬装処理が、データ加工部 6 にて行われているので、静止画像の表示が短時間で終了することはなく、音声を再生している間、継続して静止画像の表示が行われ

るようになる。

#### 【0 0 2 6】

この様にして、本発明では、静止画像を動画としてみなしてエンコード処理し、かつその長さを擬装したうえで動画ファイルに多重化することにより、動画ファイルの再生機能を有する再生装置 1 1 において、特別な機能の追加を何等なすことなく、静止画像と音声との同時再生が実現可能となる。

#### 【0 0 2 7】

ここで、図 3 を参照すると、図 3 ( a ) は、一般的な動画ファイルの形式について図示したものである。図に示すように、動画ファイルは、基本的に、映像部分である映像データと、音声部分である音声データとから構成される。更に、映像と音声との再生速度を調整し、記録時と同じタイミングで映像と音声を再生させるための情報を含んだ多重化情報（ヘッダ）を持つ。このヘッダには、音声の再生時間を示す再生時間情報も含まれている。更にはまた、多重化の方式によっては、映像と音声だけではなく、文字情報などの他の情報を含めることができる場合があり、その例としては、テロップ表示のためのテキスト情報や、各種コマンド情報等である。

#### 【0 0 2 8】

なお、一般的な多重化方式としては、MP 4 (MPEG-4System) , A S F (Advanced Streaming Format ) などがある。

#### 【0 0 2 9】

映像データは、通常は何らかのエンコード方式によってエンコードされている。これは、映像は各フレーム（静止画とみなせる）の情報を直接格納すると、サイズが大きく扱いにくいためである。さらに、一般的なエンコード方式では、単純に各フレームの情報を圧縮するだけではなく、映像固有の性質を使用したエンコード方法が用いられる。一般的な映像のエンコード方式について図 3 ( b ) に示す。

#### 【0 0 3 0】

一般的なエンコード方式では、映像の隣接したフレーム間では差分が少ないという性質を使用したエンコード方式が用いられる。映像を構成する各フレームの

うち、一番最初のフレーム及び場面転換した直後のフレームについては静止画と同等の圧縮が行われる。この静止画は I フレームと称され（通常、I は intra-coded の略である）。この I フレームはデータ量としては大きくなるが、それ単独での再生が可能である。

#### 【 0 0 3 1 】

直前のフレームとの差分が小さい場合は、直前のフレームとの差分を検出したうえで圧縮が行われる。この差分は P フレームと称される（通常、P は predictive-coded の略である）。この P フレームには、直前のフレームと差がある部分だけが情報として含まれるため、圧縮後のサイズがきわめて小さくなるが、再生の際には、直前のフレームの情報を必要とする。

#### 【 0 0 3 2 】

さらに一部のエンコード方式では、前後のフレームから差分をとることで、更に情報量を削減したうえで圧縮を行う場合がある。これを B フレームと呼ぶ（通常、B は bidirectionally-predictive-coded の略）。この B フレームは P フレームよりさらにサイズを小さくすることができる反面、再生時に、前後のフレームの情報を必要とする。また、再生時に、前後のフレームの情報を必要とするために、圧縮時にフレームの並べ替えが行われる。

#### 【 0 0 3 3 】

前後のフレームとの差分の取り方、各フレームの圧縮方式については、エンコード方式によって異なる。通常は、I フレームは極くわずかに挿入され、残りのほとんどは P フレームまたは B フレームとなるものがほとんどである。一般的なエンコード方式としては、MPEG-2, MPEG-4 Video, H. 263（IP 電話に関する ITU T 勧告）などがある。

#### 【 0 0 3 4 】

音声データについても、映像データと同様に、通常は何らかのエンコード方式によってエンコードされている。音声の場合、一定時間分の音声情報をフレームとして扱い、このフレーム単位で圧縮する方法が採られる。フレームの長さについては、全てのフレームを同じ長さとする方式もあれば、毎回フレームの長さが異なる方式も存在するが、フレームの平均的な長さとしては数十ミリ秒前後とな



る方式が多い。一般的なエンコード方式としては、m p 3 (MPEG Audio Layer I II) , A A C (Advanced Audio Coding) , WMA (Windows(R) Media Audio) などがある。

#### 【 0 0 3 5 】

本発明では、映像のエンコード方式として上記の I フレーム、P フレームを使用した圧縮を行うものを前提とする（但し、B フレームの有無はどちらでもよいものとする）。

#### 【 0 0 3 6 】

以下に、本発明の実施例の動作について詳細に説明する。前述したように、図 1 において、データ加工部 6 以外は、一般的によく知られた構成及び動作となるために、それらの詳細な説明は省略する。データ加工部 6 の動作につき説明する。図 4 はこのデータ加工部 6 により生成される、再生時間が擬装された映像データの構成例を示しており、( a ) は通常の映像データの例であり、( b ) は静止画像ファイル 2 の例である。

#### 【 0 0 3 7 】

本発明においては、データ加工部 6 における映像データの再生時間擬装方式の例として、動画エンコーダ 4 におけるエンコード方式に対応して存在するものであり、動画エンコード方式 # 1 ~ # 3 の夫々に対応して、3 種の再生時間擬装方式が採用される。よって、( c ) はエンコード方式 # 1 に対応した再生時間擬装方式による映像データの構成例であり、( d ) はエンコード方式 # 2 に対応した再生時間擬装方式による映像データの構成例であり、( e ) はエンコード方式 # 3 に対応した再生時間擬装方式による映像データの構成例である。

#### 【 0 0 3 8 】

図 5 は、図 4 ( c ) に示したエンコード方式 # 1 に対応した再生時間擬装方式による映像データの構成を生成するデータ加工部 6 の動作を示すフローチャートである。まず、入力された映像データ（動画エンコーダ 4 の出力）から、先頭の I フレームが切出される（ステップ S 1）。このとき、入力された映像データに余分な情報（時間長に関する情報など）が含まれている場合には、これが取り除かれるものとする。なお、動画エンコーダ 4 の出力が純粋に I フレームのみであ

れば、ステップ S 1 は省略されることになる。本例の場合がそれに相当する（図 4（b）参照）。

#### 【0 0 3 9】

エンコード方式 # 1 に対応した再生時間擬装方式では、無効 P フレームの連続挿入を行う。この場合のエンコード方式 # 1 は、フレームの挿入間隔（挿入周期）が制限されており、一定周期で所定フレームを挿入することが要求されるようなエンコード方式であるものとする。

#### 【0 0 4 0】

先ず、ステップ S 1 で切出された I フレームが出力され（ステップ S 2）、次に、出力されたデータの再生時間が、目的とする再生時間に達するまでの間（ステップ S 3）、一定周期（間隔：例えば 1 5 秒間隔）で、無効 P フレームの出力が行われる（ステップ S 4）。

#### 【0 0 4 1】

ここで、無効 P フレームとは、実質的な映像データを持たない P フレームを指すものとする。静止画像をエンコードしているため、先頭の I フレームに続くフレームでは、フレーム間の差分が存在しない。従って、差分が無いことを示す P フレームを挿入する。この P フレームは映像の内容に関わらず同一の内容となるため、動画エンコーダ 4 にて生成する必要はなく、予めメモリなどに保持してあるデータを出力する方法が望ましい。

#### 【0 0 4 2】

なお、MPEG-4 Video などのエンコード方式によっては、差分が無い無効 P フレームを特別な短いデータで表現することができる手段を持っていることがある（MPEG-4 Video の場合は、Not Coded フレーム）。このような手段を持っているエンコード方式では、これを使用することが望ましい。

#### 【0 0 4 3】

また、目的とする再生時間は、音声再生に必要な時間であり、この音声再生時間情報は、音声エンコーダ 5 より供給されるものとする。目的とする再生時間が終了すると、処理終了となる。なお、出力するデータが無い区間では、何も出力

されない。この方式により出力される再生時間が擬装された映像データの例は、図 4 (c) である。

#### 【0 0 4 4】

図 6 は、図 4 (d) に示したエンコード方式 # 2 に対応した再生時間擬装方式による映像データの構成を生成するデータ加工部 6 の動作を示すフローチャートである。まず、入力された映像データ（動画エンコーダ 4 の出力）から、先頭の I フレームが切出される（ステップ S 1）。このとき、入力された映像データに余分な情報（時間長に関する情報など）が含まれている場合には、これを取り除かれるものとする。なお、動画エンコーダ 4 の出力が純粋に I フレームのみであれば、ステップ S 1 は省略されることになる。本例の場合がそれに相当する（図 4 (b) 参照）。

#### 【0 0 4 5】

次に、この方式では、末尾にのみ無効 P フレームの挿入を行う。フレームの挿入間隔に制限がないエンコード方式の場合は、上記の図 4 (c) 及び図 5 により説明した方式（エンコード方式 # 1 対応の再生時間擬装方式）に代えて、この方式を用いることが望ましい。この方式では、ステップ S 1 切出された I フレームを出力した後（ステップ S 2）、末尾に無効 P フレームの挿入を行う（ステップ S 5）。この方式により出力される、再生時間擬装映像データの例が、図 4 (d) となる。

#### 【0 0 4 6】

図 7 はエンコード方式 # 3 に対応した再生時間擬装方式による映像データの構成を生成するデータ加工部 6 の動作を示すフローチャートである。この再生時間擬装方式では、無効 P フレームによって再生時間を延長するのではなく、再生時間を情報 T として持たせることで延長を行う。再生時間は映像データの内部または多重化情報に埋め込まれる。この方式を使用できるのは、再生装置が映像データまたは多重化情報に埋め込まれた再生時間情報 T を確実に参照する保証がある場合のみである。

#### 【0 0 4 7】

この方式では、ステップ S 1 で切出された I フレームを出力した後（ステップ

S 2)、再生時間情報 T が出力される (ステップ S 6)。このステップ S 6 の実現方法は、再生時間情報 T が埋め込まれる場所によって異なり、映像データに埋め込まれる場合は、ステップ S 2 で出力されたデータの改変を行い、多重化情報に埋め込まれる場合は、マルチプレクサ 7 に再生時間情報 T を伝達する。この方式により出力される再生時間擬装映像データの例が、図 4 (e) に示されるものである。

#### 【 0 0 4 8 】

記録装置 1 側では、図 4 (c) ~ (e) に示した再生時間擬装の各方式により得られた動画ファイルが出力されるので、再生装置 1 1 側では、これ等再生時間擬装の各方式に対応した動画エンコード方式 # 1 ~ # 3 にそれぞれ対応して、動画デコーダ 1 3 を設けておく必要があることは明白である。

#### 【 0 0 4 9 】

図 8 は動画エンコード方式 # 1 ~ # 3 の全てに対応した再生時間擬装方式を示すデータ加工部 6 の動作を示すフローチャートである。図 5 ~ 図 7 と同等ステップは同一符号により示されている。この場合、ステップ S 1 における処理 (先頭の I フレームを切出す処理) の後に、動画エンコーダ 4 におけるエンコード方式 # 1 ~ # 3 に応じて、上述した図 5 ~ 図 7 のフローチャートに従った再生時擬装方式を択一的に選択するものであり、そのために、ステップ S 7 が追加されている。このステップ S 7 の後に、各方式 # 1 ~ # 3 に対応した処理 (図 5 ~ 図 7 に示した S 3 以降の処理) が行われることになる。

#### 【 0 0 5 0 】

このステップ S 7 による方式選択の方法としては、送信者側から指定 (手動や自動) する方法、送信相手 (アドレス) に応じて指定する方法、静止画の種類に応じて手動や自動にて行う方法がある。静止画の種類に応じて選択する手法の例として、図 9 に示す如く定めておくことができる。例えば、静止画の種類が写真であれば、エンコード方式 # 1 として M P E G - 2、絵であれば、エンコード方式 # 2 として M P E G - 4、名刺や手紙、メモ等の文字であれば、エンコード方式 # 3 として H. 2 6 3 とすることができるが、これに限定されるものではない。

## 【0051】

なお、送信側である記録装置1において、エンコード方式#1～#3のいずれかを特定するための情報を、動画ファイルのヘッダ部に付加し、受信側の再生装置11における動画デコーダ13でこれを読取ってエンコード方式を判断し、各方式を選択するよう構成する。

## 【0052】

以上述べた如く、上記実施例では、静止画像と音声とを動画ファイルとして作成するために、再生装置において、静止画と音声との同時再生機能を特別に追加することなく、当該機能を実現できることになる。よって、再生装置は一般の動画再生装置がそのまま使用可能である。

## 【0053】

また、記録装置において、動画の記録機能に対してデータ加工部6の機能を追加するだけで、本発明が実現できるので、音声付き静止画再生処理システムの実現が極めて容易である。動画エンコーダや音声エンコーダ、マルチプレクサは、動画を記録する機能で用いられるものと全く同一のものが使用可能である。

## 【0054】

また、本発明により作成される動画ファイルのサイズ（容量）は、基本的にはIフレーム（静止画1フレーム）分のサイズと音声データのサイズとの合計とほぼ一致するので、少ないデータサイズとなる。これは、静止画像を映像としてエンコードする際に、画像が時間的に変化しないことを利用して、極力、余分なデータの発生を抑えているためであり、よって通信データ量も少くなる。

## 【0055】

図10は本発明の他の実施例を示す記録装置1の機能ブロック図であり、図1と同等部分は同一符号にて示している。図10においては、カメラ付き携帯端末（携帯電話機に代表されるもの）に、本発明を適用したものであり、図1の静止画像ファイル2と音声ファイル3の代りに、カメラ8とマイク9とを設けたものである。

## 【0056】

カメラ8により撮影された画像とマイク9により集音された音声とを、直接動

画エンコーダ 4 と音声エンコーダ 5 とへそれぞれ供給することによって、一般の動画撮影の場合と同様に、その場の映像（静止画）と音声とを動画ファイルとして生成するものである。これにより、カメラによる静止画撮影時に、同時に音声を記録することができ、なおかつ生成された動画ファイルを一般の動画再生装置にて再生することによって、静止画と音声の同時再生が可能になる。

#### 【0057】

本発明の更に他の実施例を図 11 に示す。この実施例では、記録装置 1 において、静止画像ファイルと音声ファイルとを保持する代わりに、デマルチプレクサ 21、動画デコーダ 22、音声デコーダ 23 による再生装置の機能を追加し、任意の動画ファイルを変換して、静止画とそれに関連する音声とを抽出する。この実施例では、動画デコーダ 22 の直後にフレーム抽出部 24 を設けており、映像に関して動画エンコーダ 22 の出力から 1 フレームを静止画として切出す。この場合、切出すフレームは、先頭のフレーム、最後のフレーム、中央のフレームなど、所定のフレームが考えられる。これにより、任意の動画ファイルから静止画と音声とを抽出して、新たに静止画と音声とを格納した動画ファイルを作成することが可能になる。

#### 【0058】

すなわち、図 11 の例では、デマルチプレクサ 21、動画デコーダ 22、音声デコーダ 23、フレーム抽出部 24、（音声切出し部 25）により、動画－静止画変換を行って、任意の動画ファイルから所望の静止画を切出し、この切出された音声付き静止画を、先の実施例と同様の処理により、動画ファイルを生成するようにしている。

#### 【0059】

また、図 11 では点線で示しているが、音声に関して音声切出し部 25 を設け、音声デコーダ 23 の出力のうち任意の部分を切出した上で、音声エンコーダ 5 に入力する処理を追加することによって、任意の動画ファイルの任意の位置から静止画と音声とを切出し、これを単一の動画ファイルとして構成することが可能になる。一般的な音声エンコード方式の場合、音声の切出しは単純なフレーム単位での切出し処理で問題ないため、音声切出し部 25 は単純なデータ加工となる

。

**【 0 0 6 0 】**

図 1 2 は本発明の別の実施例を示す動作フローチャートであり、図 8 と同等ステップは同一符号にて示している。この実施例は、データ加工部 6 の機能を変更して、一つの動画ファイル内に複数の静止画を埋め込むことを可能にしている。すなわち、本実施例では、複数の静止画に説明の音声が付加するものであり、いわゆる紙芝居方式のファイルを作成する場合のものである。

**【 0 0 6 1 】**

この実施例により作成される映像データの例が図 1 3 に示されている。この実施例では、図 1 の記録装置 1 内のデータ加工部 6 の動作であるフローチャートに、図 1 2 に示す如くステップ S 8 を追加し、I フレームの埋め込みを複数回実行するように変更している。更に、記録装置内の静止画像ファイル 2 を複数枚用意し、これらを動画エンコーダ 4 にて全て I フレームにエンコードした後、データ加工部 6 にて、順番に I フレームを映像データに取込む。

**【 0 0 6 2 】**

こうして生成される映像データは、図 1 3 に示すように、1 つの映像データの中に複数の I フレームが埋め込まれており、図 1 の再生装置 1 1 において、静止画像ファイルの内容を、紙芝居的に順次表示させることができる。用意する静止画像ファイルは、カメラの連写機能などを用いて連続撮影した画像を用いることが望ましい。

**【 0 0 6 3 】**

なお、図 5 ～図 7 のフローチャートの夫々の最後に、図 1 2 に示したステップ S 8 を追加することもできることは明白である。

**【 0 0 6 4 】**

上記各実施例におけるデータ加工部 6 の動作は、予めその動作手順をプログラムとして記録媒体に格納しておき、これをコンピュータ（C P U）により読取って実行させるようにすることができるものである。

**【 0 0 6 5 】****【発明の効果】**

本発明によれば、所定の動画像符号化方式を用いて符号化された静止画像データに対して、動画再生装置において静止画像の再生を所定時間継続させるための情報を当該動画像符号化方式に応じた形式で付与してファイルを作成するようにしているので、このファイルを動画再生装置で再生する場合、符号化された静止画像データに付与された情報により、静止画像の表示を所定時間継続させることができる。その結果、このファイルとこれに対応する音声データとを多重化することにより得られた動画ファイルを、動画再生装置で再生する場合、音声を再生している間、静止画像の表示を継続して行うことが可能となる。

#### 【0066】

すなわち、再生装置側では、特別の静止画再生機能を付加することなく、音声付きの静止画を再生することができ、また、データ量も、基本的に、1フレーム分のデータと音声データとの和となるために、極めて少なくなり、記録装置側から再生装置側への送信データ量が少なくなり、通信網の輻輳もなくなり、また再生装置側での処理も高速化が図れることになる。

#### 【図面の簡単な説明】

##### 【図1】

本発明の一実施例による記録装置及び再生装置の構成を示す図である。

##### 【図2】

本発明の一実施例による動画ファイル作成の過程を示す図である。

##### 【図3】

一般的な動画ファイルの形式及び一般的なエンコード方式について説明するための図である。

##### 【図4】

図1のデータ加工部6によって生成される再生時間擬装映像データの構造の例を示す図である。

##### 【図5】

本発明のデータ加工部6の動作の一例を示すフローチャートである。

##### 【図6】

本発明のデータ加工部6の動作の他の例を示すフローチャートである。



**【図 7】**

本発明のデータ加工部 6 の動作の更に他の例を示すフローチャートである。

**【図 8】**

本発明のデータ加工部 6 の動作の別の例を示すフローチャートである。

**【図 9】**

静止画の種類と動画エンコード方式との対応例を示す図である。

**【図 1 0】**

本発明の他の実施例による記憶装置の構成を示す図である。

**【図 1 1】**

本発明の更に他の実施例による記憶装置の構成を示す図である。

**【図 1 2】**

図 1 1 の記憶装置 1 のデータ加工部 6 の動作例を示すフローチャートである。

**【図 1 3】**

図 1 2 の動作例に従って生成される再生時間擬装映像データの構造の例を示す図である。

**【符号の説明】**

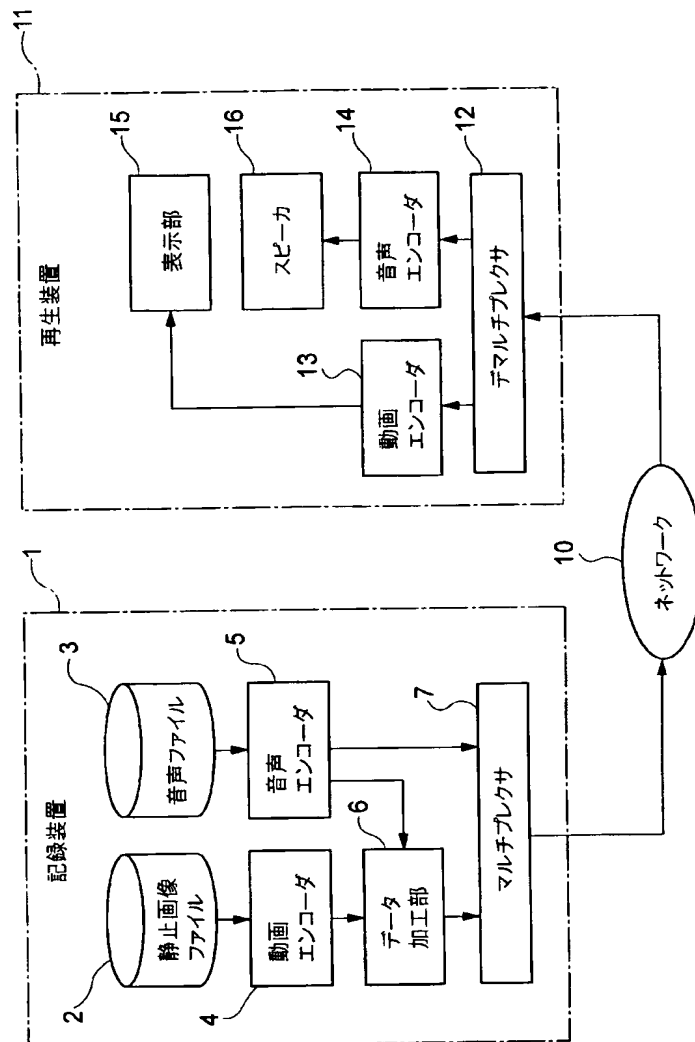
- 1 記憶装置
- 2 静止画像ファイル
- 3 音声ファイル
- 4 動画エンコーダ
- 5 音声エンコーダ
- 6 データ加工部
- 7 マルチプレクサ
- 8 カメラ
- 9 マイク
- 1 0 ネットワーク
- 1 1 再生装置
- 1 2, 2 1 デマルチプレクサ
- 1 3, 2 2 動画デコーダ

- 1 4, 2 3 音声デコーダ
- 1 5 表示装置
- 1 6 スピーカ
- 2 4 フレーム抽出部
- 2 5 音声切出し部

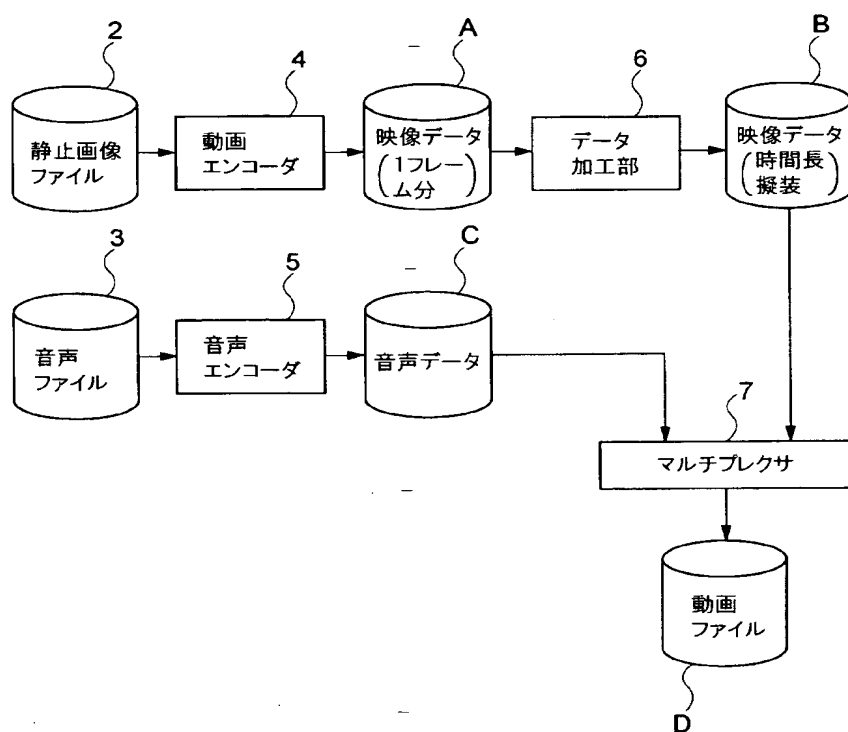
【書類名】

図面

【図 1】

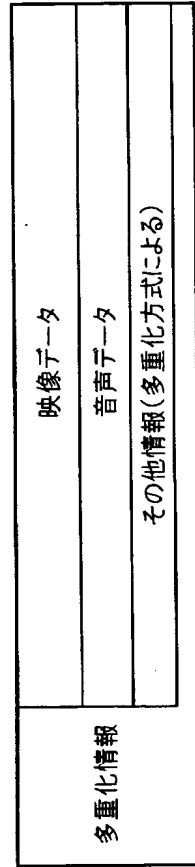


【図 2】

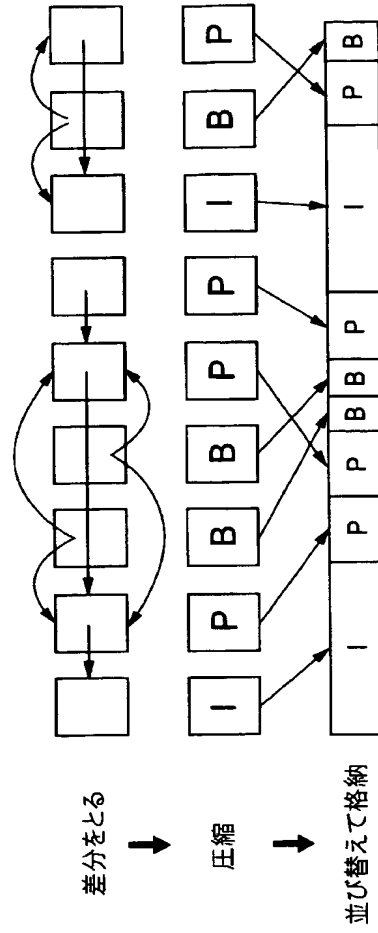


【図 3】

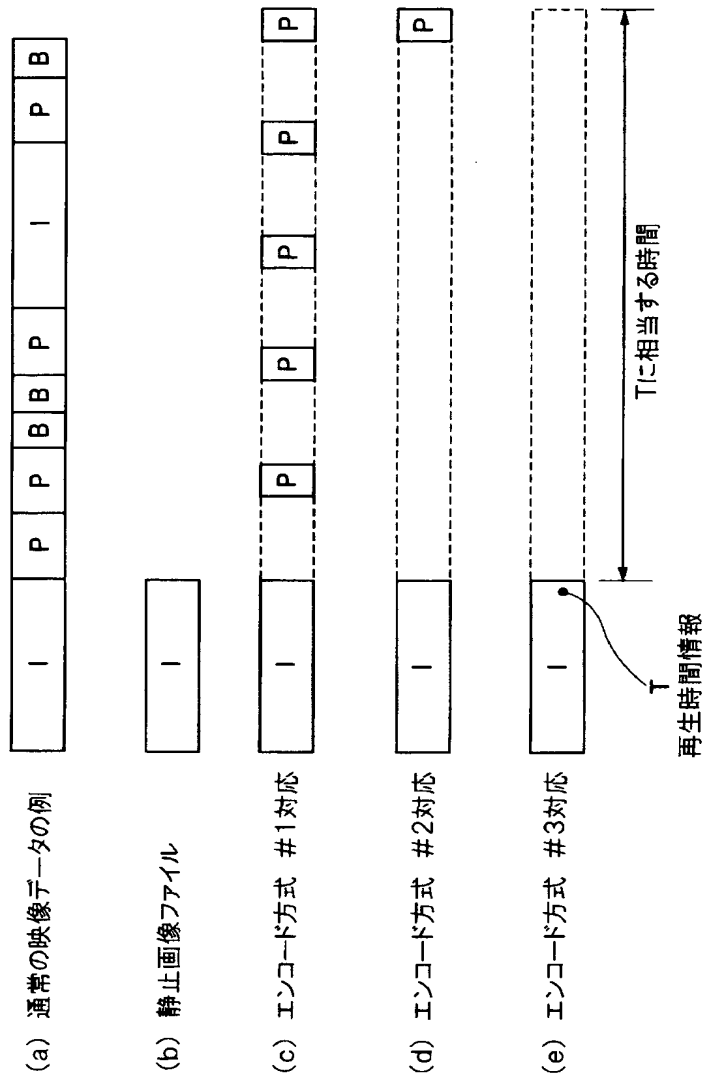
(a) 動画ファイルの形式



(b) 映像のエンコード方式

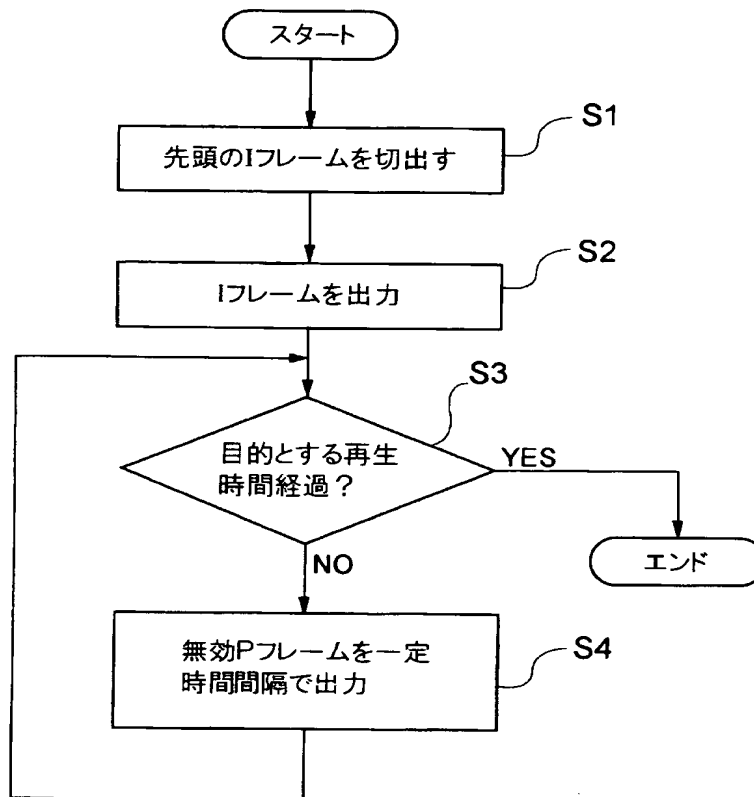


【図 4】



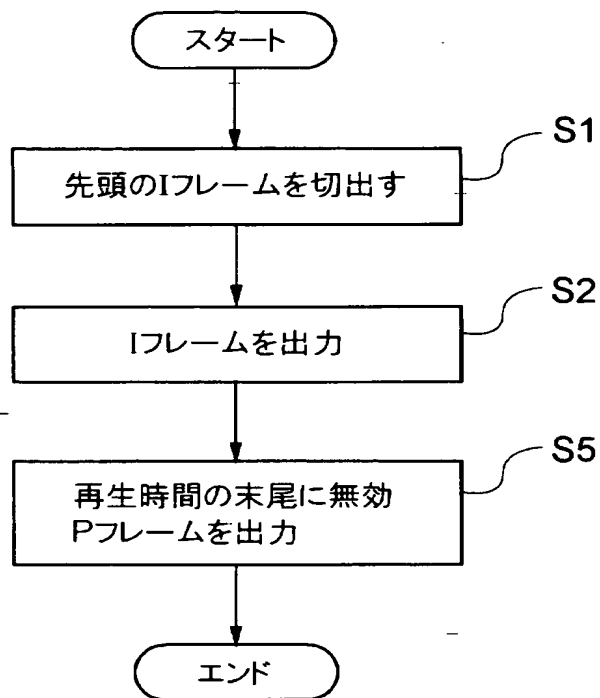
【図 5】

エンコード方式#1対応のフローチャート



【図 6】

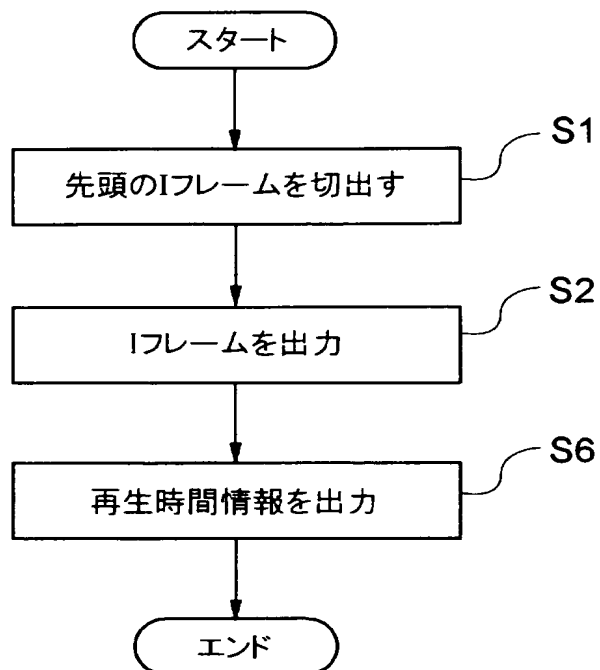
エンコード方式#2対応のフローチャート





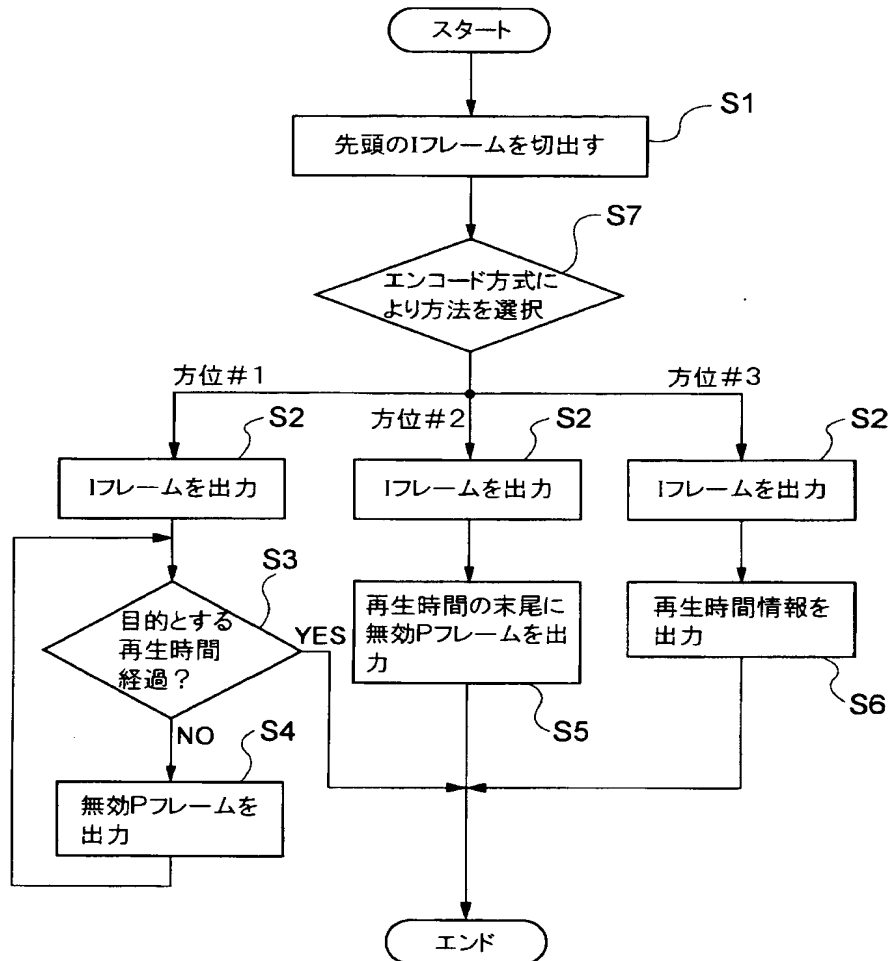
【図 7】

エンコード方式#3対応のフローチャート



【図 8】

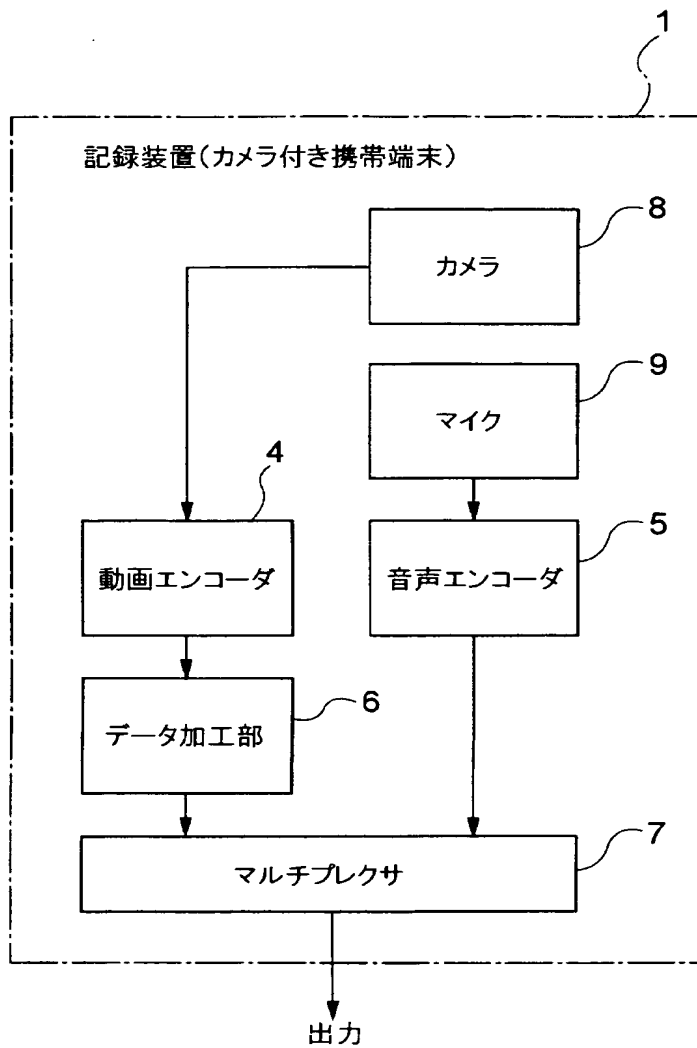
エンコード方式 #1～#3 全てに対応のフローチャート



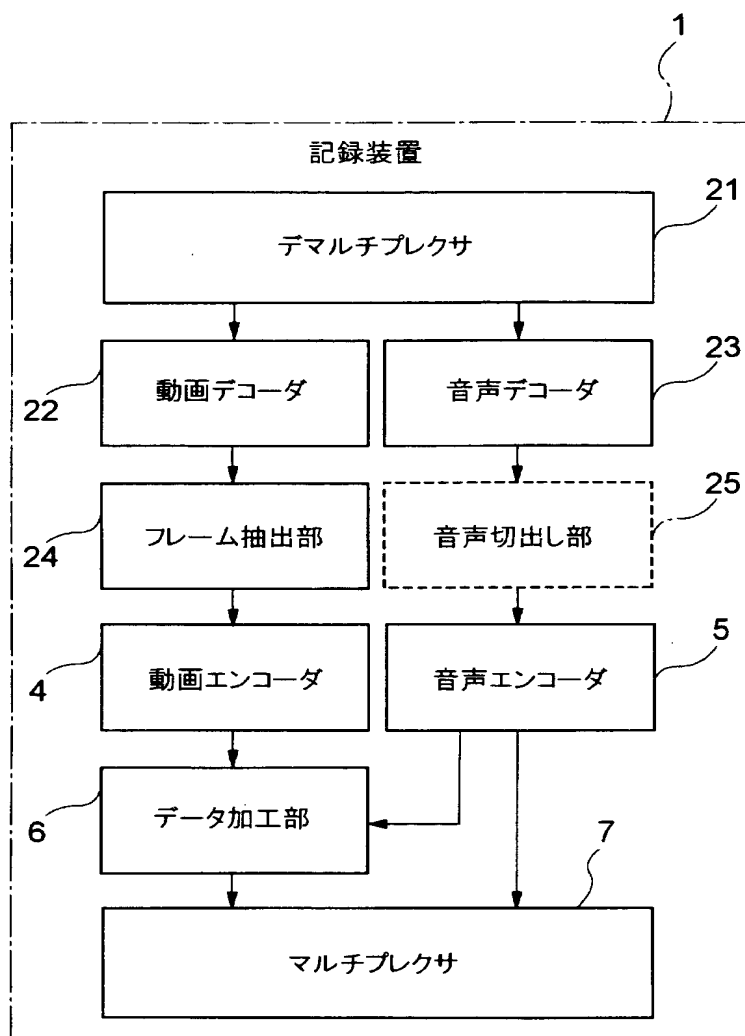
【図 9】

静止画の種類	エンコード方式 #	エンコード方式
写真	# 1	MPEG-2
絵	# 2	MPEG-4
文字	# 3	H. 263

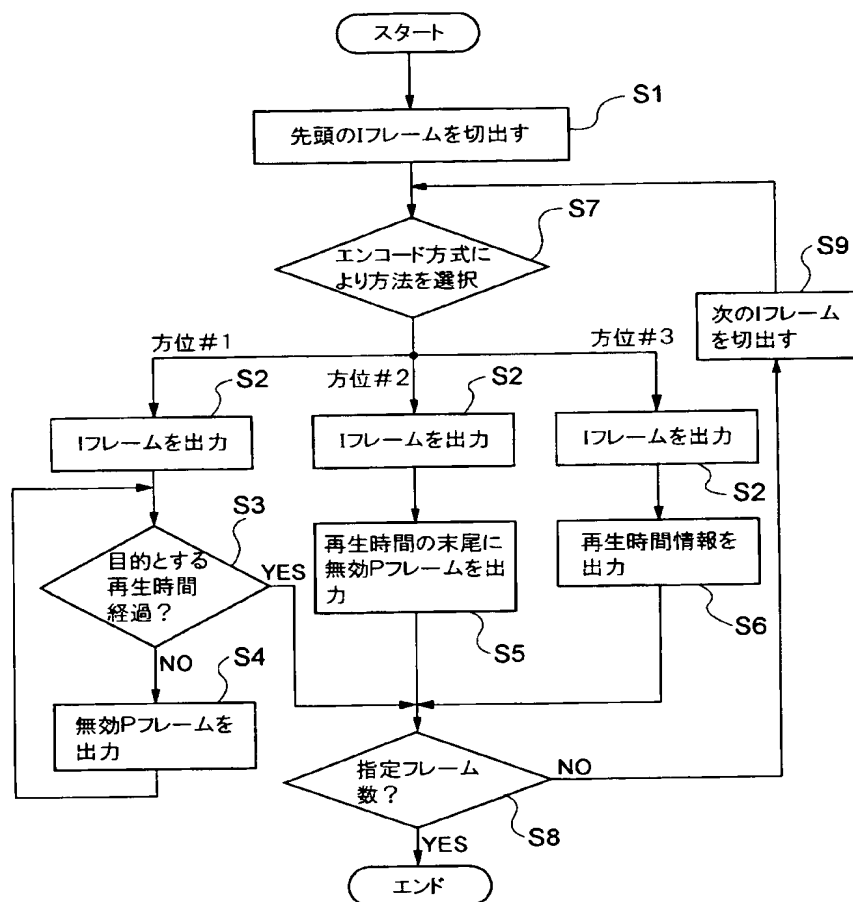
【図 10】



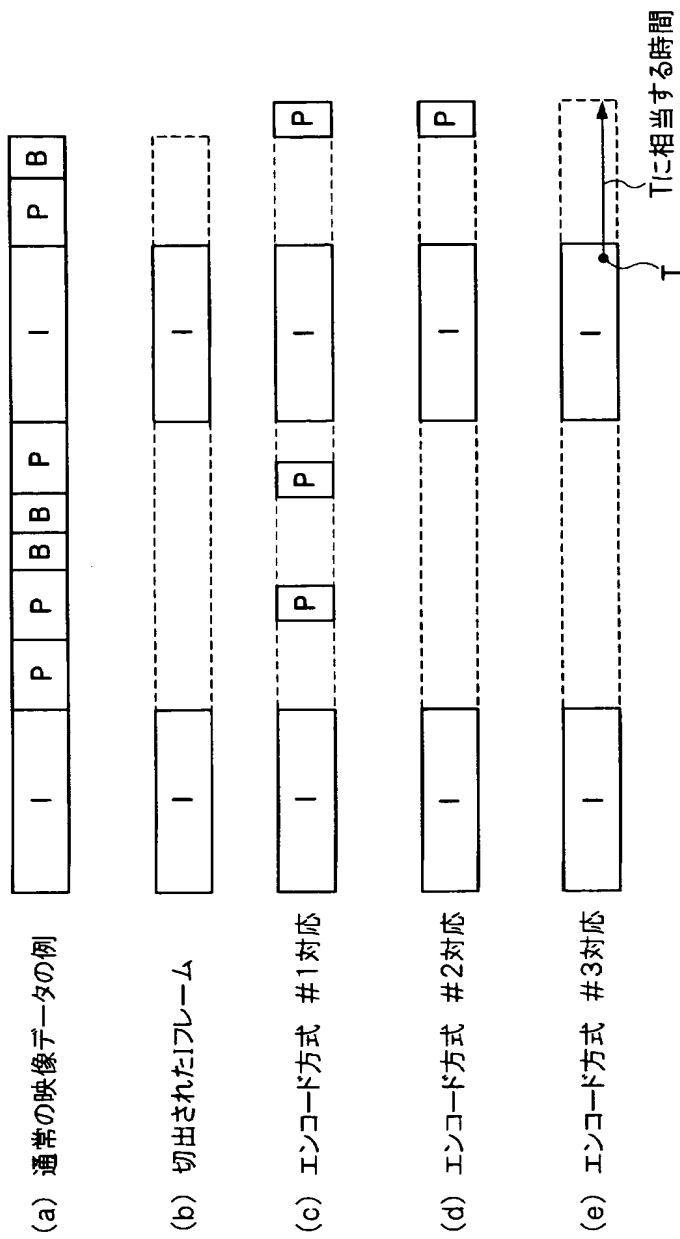
【図 11】



【図 12】



【図 13】



【書類名】 要約書

【要約】

【課題】 特定の形式や専用の機能をもたない再生装置でも、静止画と音声の同時再生が可能となるよう、生成するファイルは一般的な動画ファイルの形式を守りつつ静止画と音声を単一ファイルとして作成することができるようにする。

【解決手段】 動画エンコーダ 4、音声エンコーダ 5 及びマルチプレクサ 7 を有する記録装置 1 において、動画エンコーダ 4 及びマルチプレクサ 7 間に、データ加工部 6 を設ける。データ加工部 6 は、動画エンコーダ 4 により符号化された静止画像データに、動画再生装置 1 1 において静止画像の再生を所定時間継続させるための情報を動画エンコーダ 4 の動画像符号化方式に応じた形式で付与して出力する。

【選択図】 図 1



特願 2 0 0 3 - 0 8 4 0 3 4

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 4 2 3 7 ]

1. 変更年月日

1 9 9 0 年 8 月 2 9 日

[変更理由]

新規登録

住 所

東京都港区芝五丁目 7 番 1 号

氏 名

日本電気株式会社